DataSwap

Enhancing Collaborative Machine Learning

Background

The Data Problem

The importance of GPUs to AI is well-known, as is their scarcity. However, the need for high-quality data, crucial for AI, is often overlooked. Sites like Twitter and Reddit provide valuable data that is becoming harder and more expensive to access.

Data Accessibility and Costs

Twitter's decision to charge $5,000 monthly for API access to a million tweets highlights a significant hurdle for AI startups: accessing affordable, high-quality data. This issue, alongside the GPU shortage, is a major barrier for new entrants.

The AI Space Monopoly

The dominance of major companies over crucial AI training data is leading to a monopolized AI industry. This control limits competition and innovation, with a few big players like OpenAI shaping the future of AI.

DataSwap

DataSwap is a first-of-its-kind marketplace, tailored specifically for the AI industry. It provides a secure and anonymous platform for companies to either exchange their existing training datasets or find new, high-quality datasets. DataSwap aims to streamline the data acquisition and sharing process, making it simpler and more cost-effective for AI developers and researchers to access the precise data they need for training cutting-edge AI models.

SWAP

SWAP is a unique protocol combining cryptography and reputation systems to allow AI startups to safely swap datasets.

How does SWAP work?

1. Submit Sample Data.
2. Specify Desired Data.
3. Find and acquire the desired data in the marketplace.
4. Deposit $SWAP tokens as collateral.
5. Execute the data swap.
6. Define a return period.
7. Return data after the period ends.
8. Release $SWAP collateral.

Types of Swaps:

Direct Swap: Two users exchange their datasets directly, each wanting what the other possesses.

Multiparty Swap: The system identifies and orchestrates swaps among multiple parties to ensure everyone receives their desired datasets.

Data Purchase: Users or companies lacking tradable assets can buy datasets with SWAP tokens or USD, ideal for acquiring labeled or pre trained data.

Hacker Service: For highly specific or proprietary data needs, the DS hacker service employs developer expertise to source any type of data.